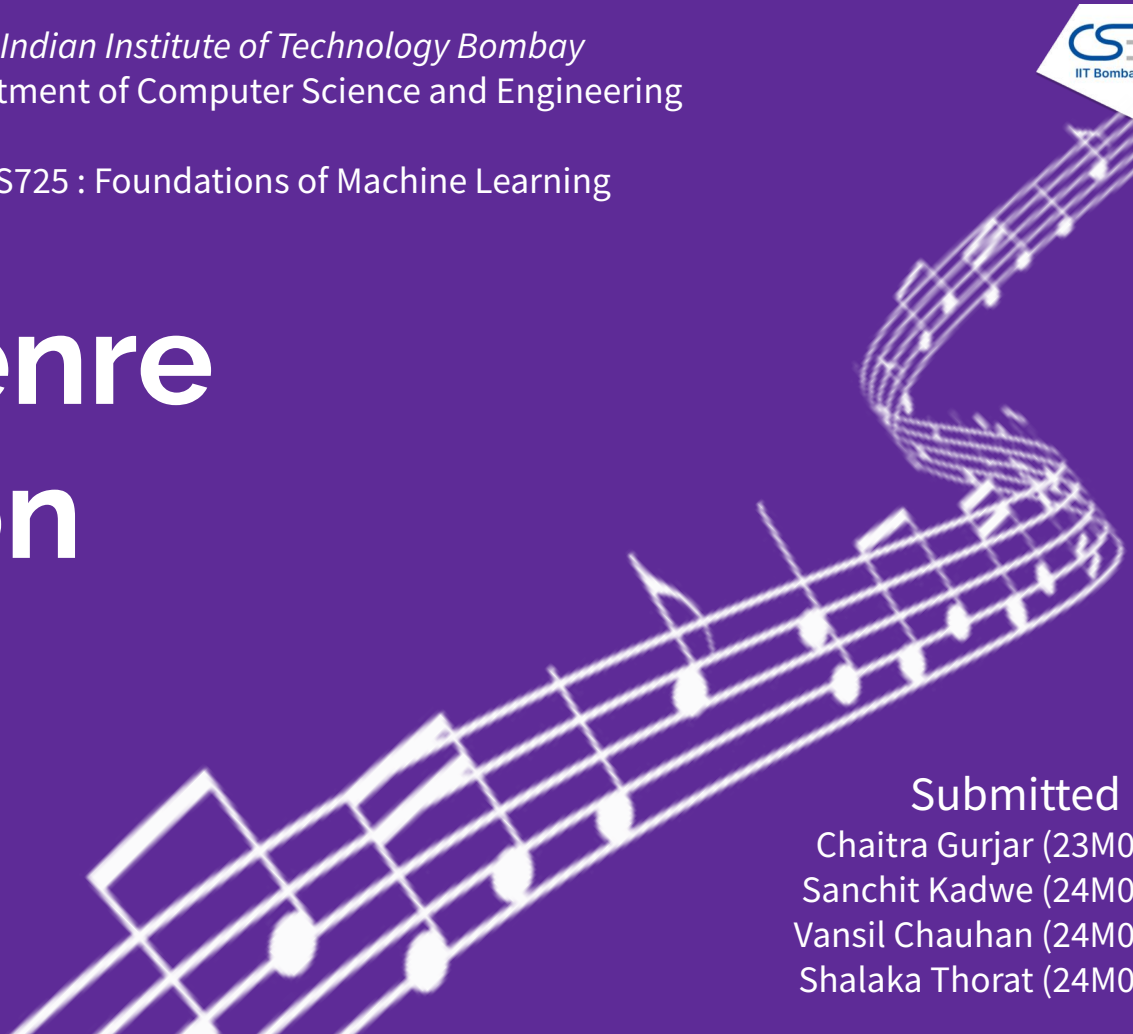
Indian Institute of Technology Bombay
Department of Computer Science and Engineering

CS725 : Foundations of Machine Learning

Music Genre Prediction

Team Musica



Submitted By :
Chaitra Gurjar (23M0831)
Sanchit Kadwe (24M0836)
Vansil Chauhan (24M0847)
Shalaka Thorat (24M0848)

Submitted on :
Monday, 25th November, 2024

Problem Statement

- This project aims to classify music tracks into different genres using machine learning techniques.
- The GTZAN dataset is used, containing 10 genres with 100 tracks each, totalling 1000 tracks.
- The main objectives are to preprocess the data, extract meaningful features, train a classifier using CNN, and evaluate its performance.
- Machine learning techniques used include feature extraction, classification, CNN and evaluation metrics.
- Finally we demonstrate the project using an app which takes a .wav file as input and predicts it genre.

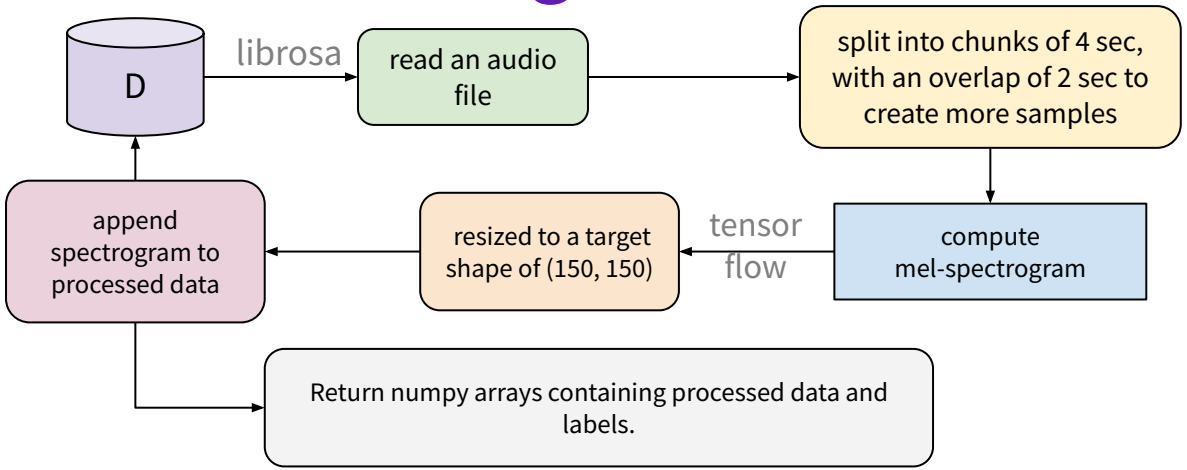
Dataset Used

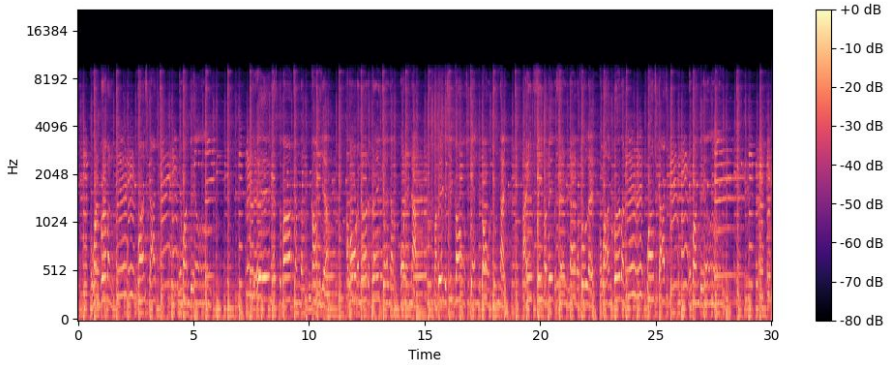
- The GTZAN dataset is a widely used benchmark in music genre classification.
- It includes 10 genres: Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae, and Rock.
- Each genre consists of 100 audio tracks, each 30 seconds long, recorded in 22050 Hz Mono 16-bit format.
- The dataset provides a balanced distribution of genres, ideal for training and evaluating classifiers.
- Data is stored in .wav format, making it suitable for audio processing and feature extraction.

Previous Attempts

- We explored csv data from 2 files which consisted of attributes derived from the audio files.
- We ran various algorithms such as Logistic Regression, SVM, KNN, Decision Trees, Random Forest etc.
- The best accuracy achieved was around 88%.
- As it couldn't improve further, so we decided to take up audio file, process mel-spectrograms, and use CNNs for predictions.

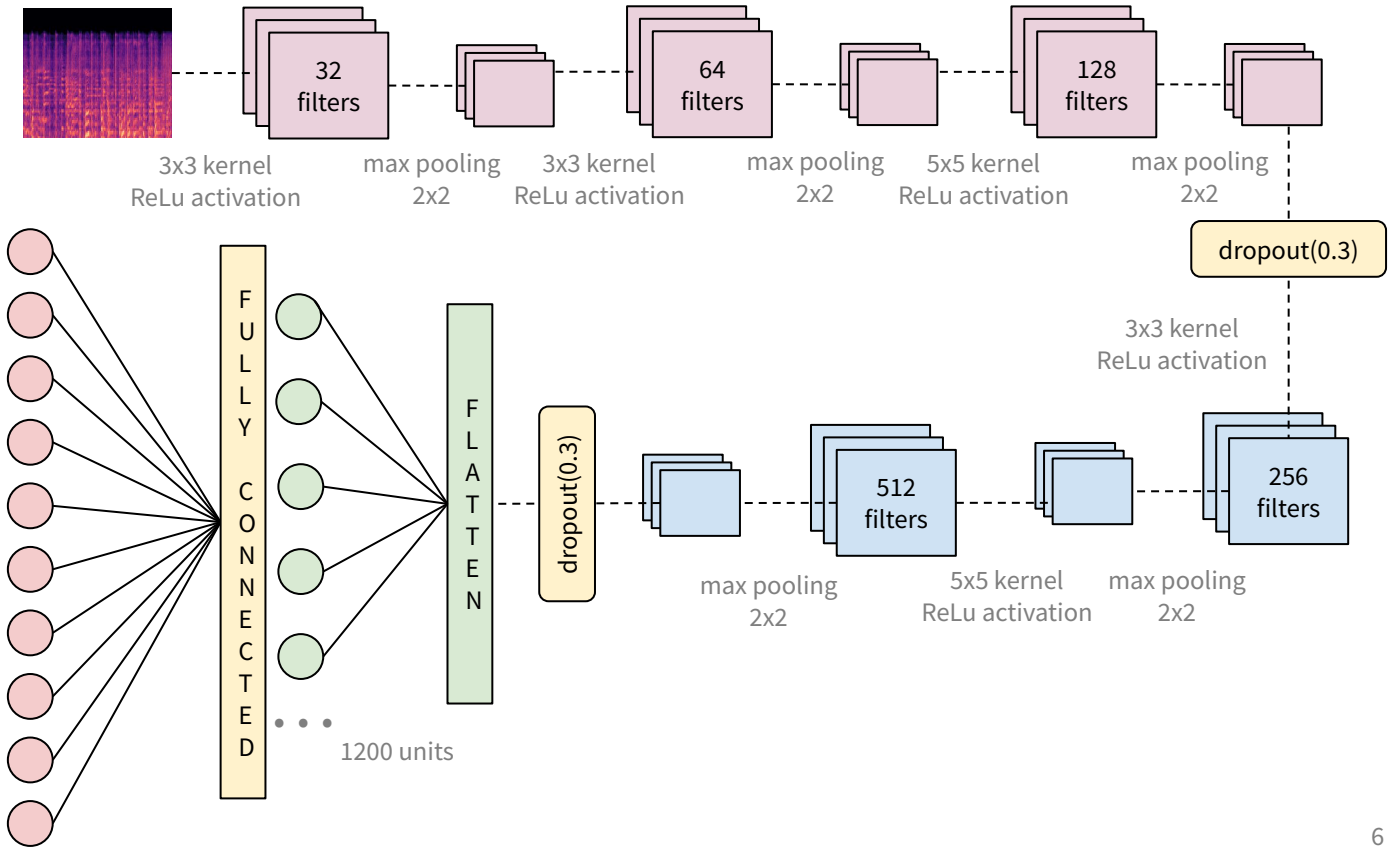
Pre-Processing the Data



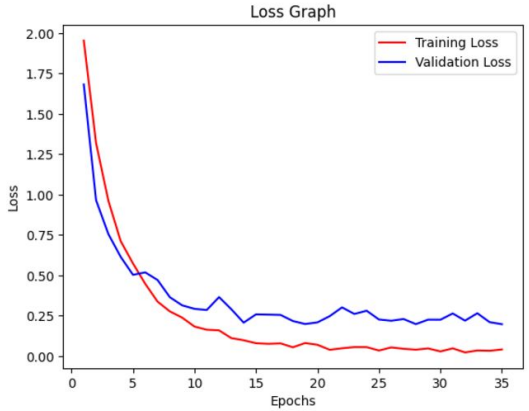


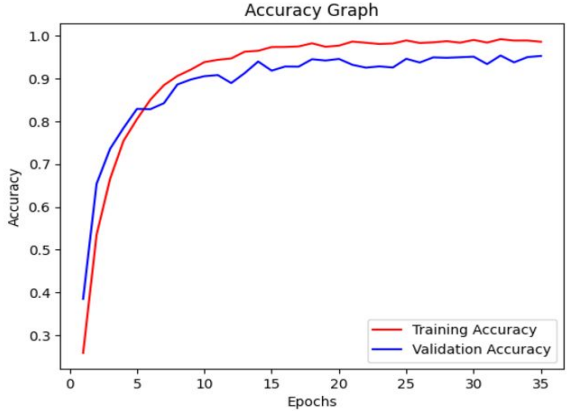
Spectrograms: Provide a time - frequency representation of the audio signals.

Model Training



Model Evaluation

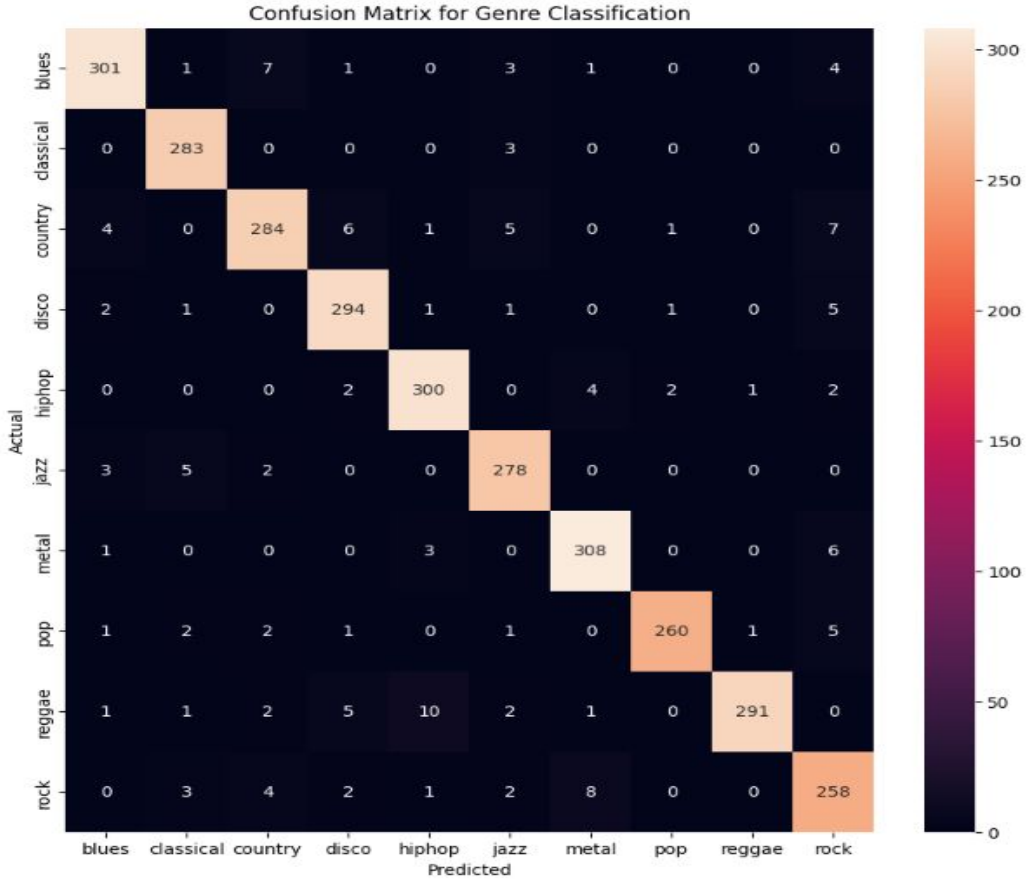




	precision	recall	f1-score	support
blues	0.96	0.95	0.95	318
classical	0.96	0.99	0.97	286
country	0.94	0.92	0.93	308
disco	0.95	0.96	0.95	305
hiphop	0.95	0.96	0.96	311
jazz	0.94	0.97	0.95	288
metal	0.96	0.97	0.96	318
pop	0.98	0.95	0.97	273
reggae	0.99	0.93	0.96	313
rock	0.90	0.93	0.91	278
accuracy			0.95	2998
macro avg	0.95	0.95	0.95	2998
weighted avg	0.95	0.95	0.95	2998

- Problem Statement
- Dataset Used
- Previous Attempts
- Pre Processing
- Model Training
- Performance
- Demonstration
- Team Contributions

Confusion Matrix



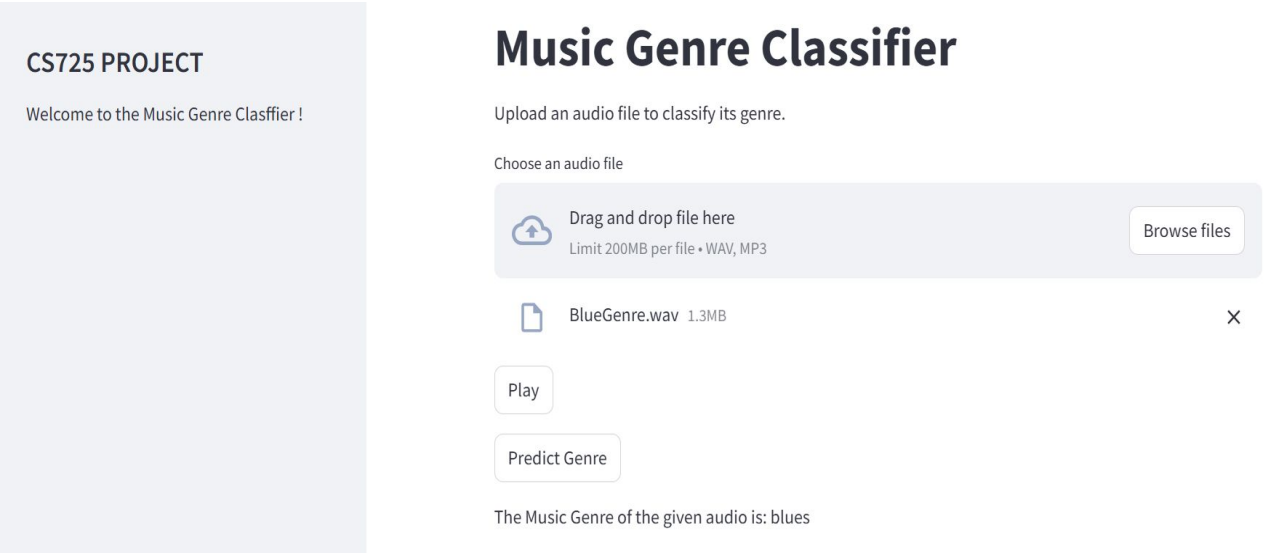
Comparison of Models

Algorithm	Using 3 sec Audio File Data	Using 30 sec Audio File Data
SVM	85.3	69
Decision Tree	64.2	46.6
Random Forest	86.1	66.6
Logistic Regression	73.3	69.3
KNN	88.7	65.3
CNN using Mel-Spectrograms	NA	95.29

- Problem Statement
- Dataset Used
- Previous Attempts
- Pre Processing
- Model Training
- Performance
- Demonstration
- Team Contributions

A Quick Demo…

https://cs725-genreclassifier.streamlit.app/



Team Contributions

- Chaitra & Shalaka:
  - Chaitra : Literature Survey, CSV Data Exploration, Analysis
  - Shalaka : ML Implementation for CSV Data, CNN Fine-Tuning
- Sanchit & Vansil:
  - Sanchit : Mel-Spectrogram Generation, CNN Model Training & Evaluation
  - Vansil : KNN and Logistic Reg. Implementation on CSV Data
- Team Efforts:
  - Documentation: Presentation & Codebase Updates

Future Work & References

Future Work:

- Doing Hyperparameter Tuning and coming up with a better Model than current Implementation.
- Balancing the Dataset, in order to avoid misclassifications for a specific Genre. (Rock Genre has more misclassifications currently.)
- Combining Mel-Spectrogram and CSV Features and analyzing the effect on Model Training and Evaluation.

References:

- https://www.clairvoyant.ai/blog/music-genre-classification-using-cnn
- https://paperswithcode.com/dataset/gtzan
- https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53
- https://medium.com/@namratadutt2/music-genre-classification-using-cnn-part-2-classification-ee5400cfbc4f
- https://youtu.be/KW6585XMV3c?si=AkLzthzMceVdHmuw